The anatomy of reliability: A must read for future human brain mapping

Xiu-Xia Xing[1], Xi-Nian Zuo[2,3,4,5,6*]

[1] School of Applied Sciences, Beijing University of Technology, Beijing, China; [2] Key Laboratory of Brain and Education, Nanning Normal University, Nanning, Guangxi, China; [3] Department of Psychology, University of Chinese Academy of Sciences (UCAS), Beijing, China; [4] CAS Key Laboratory of Behavioral Sciences, Institute of Psychology, Beijing, China; [5] Magnetic Resonance Imaging Research Center and Research Center for Lifespan Development of Mind and Brain (CLIMB), CAS Institute of Psychology, Beijing, China; [6] Institute for Brain Research and Rehabilitation, South China Normal University, Guangzhou, Guangdong, China

*Corresponding Author:

Xi-Nian Zuo, PhD

Professor of Psychology
Department of Psychology, Chinese Academy of Sciences (CAS)

PI, CAS Key Laboratory of Behavioral Science

Director
Magnetic Resonance Imaging Research Center,
Research Center for Lifespan Development of Mind and Brain
CAS Institute of Psychology

Address:
    Southern Building, Room 708
    NO 16 Lincui Road, Chaoyang District
    CAS Institute of Psychology
    Beijing, 100101, China

Emails:
    zuoxn@psych.ac.cn or zuoxn@gxtc.edu.cn

Phone/Fax:
    +861064853798

Main Text

Human brain mapping (HBM) is increasingly becoming a multi-disciplinary field where some scientific issues are fundamental for all scientists and applications of using the technology to investigate individual differences. Reliability represents a significant issue for all scientific fields and has particularly been overlooked for decades by the HBM field [1]. Meanwhile, recent advances in open science have offered the field big data for developing novel methodological frameworks as well as performing large-scale investigations of the brain-mind associations based upon the individual differences assessed with HBM [2]. A systematic investigation of reliability seems still far behind these HBM developments. It is critical that reliability is evaluated ahead of these applications, motivating the current commentary on delineation of the anatomy of reliability for future HBM.

Reliability has a specific meaning in probability theory and is defined as a statistic on characterizing stochastic processes of individual variability. Given an HBM measure, intra-individual difference is variability of the repeated measurements across different occasions within the individual or within the subject undergoing HBM (i.e., within-subject variability). Inter-individual difference is the variability of the measurements between different individuals or subjects (i.e., between-subject variability). The following equation mathematically defines reliability, where V_b and V_w are between-subject and within-subject variability, respectively.

$$\mathrm{ICC} = \frac{V_b - V_w}{V_b + V_w}$$

Mapping of reliability as a function of V_b and V_w generates a figure of the anatomy of reliability (see Figure 1). This anatomy map clearly indicates that reliability is a composite metric of both V_b and V_w, making itself a relative metric rather than an absolute measurement. Any location of the reliability field map describes a combination of both within-subject and between-subject variability (Figure 1a). High reliability of an HBM measure requires that the within-subject variability of the measure is relatively small comparing with its between-subject variability. This is particularly crucial for clinical applications [3], which call for a measure both stable across time for an individual (i.e., low within-subject variability) and recognizable between different individuals (i.e., high between-subject variability) [1]. Reliability is usually quantified by intra-class correlation (ICC). For the guidance on practical evaluation of the reliability, the values of ICC can be categorized into some common intervals, among which two most popular categories are: 1) $0 < \mathrm{ICC} \leq 0.2$ (slight), $0.2 < \mathrm{ICC} \leq 0.4$ (fair), $0.4 < \mathrm{ICC} \leq 0.6$ (moderate), $0.6 < \mathrm{ICC} \leq 0.8$ (substantial), $0.8 < \mathrm{ICC} \leq 1.0$ (almost perfect) [4] and 2) $0 < \mathrm{ICC} \leq 0.5$ (poor), $0.5 < \mathrm{ICC} \leq 0.75$ (moderate), $0.75 < \mathrm{ICC} \leq 0.9$ (good), $0.9 < \mathrm{ICC} \leq 1$ (excellent) [5]. Figure 1b depicts the ICC contour plot of the reliability field map with the category 2. In clinical practice, a minimal reliability of almost perfect level (≥ 0.8) is needed [6].

Open neuroscience has started to establish rich data resources for evaluation of reliability with HBM [7]. It is becoming a reality that researchers could include the reliability of the measures employed in their studies by using the reliability field map. This should become a standard for reliable HBM, especially for studies with novel methodology proposed [8,9]. We believe that reliable HBM will greatly advance clinical practice in the future [10-12].
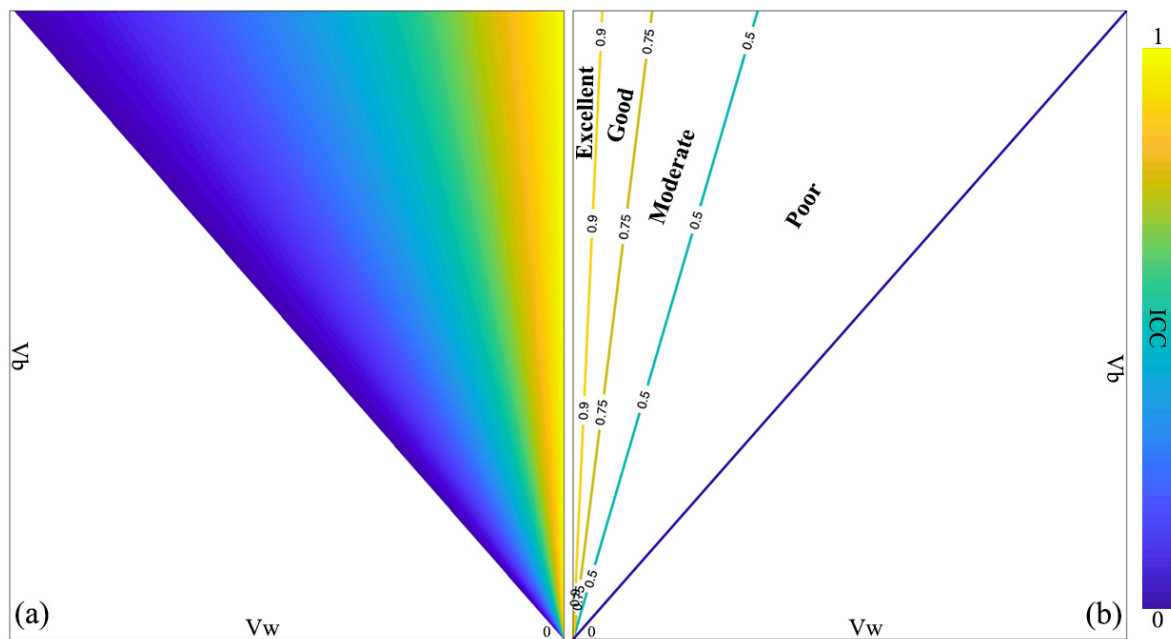
Figure 1. The anatomy of reliability. (a) The ICC field map (b) The ICC contour map

References

[1] Zuo XN, Xing XX. Test-retest reliabilities of resting-state FMRI measurements in human brain functional connectomics: a systems neuroscience perspective. Neurosci Biobehav Rev 2014; 45:100-18.

[2] Van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K; WU-Minn HCP Consortium. The WU-Minn Human Connectome Project: an overview. Neuroimage 2013; 80:62-79.

[3] Kraemer HC. The reliability of clinical diagnoses: state of the art. Annu Rev Clin Psychol 2014; 10:111-30.

[4] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33(1):159-74.

[5] Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med 2016; 15(2):155-63.

[6] Joseph LF, Bruce L, Myunghee CP. Statistical Methods for Rates and Proportions. 3rd ed. New Jersey: Wiley, 2003.

[7] Zuo XN, Anderson JS, Bellec P, Birn RM, Biswal BB, Blautzik J, Breitner JC, Buckner RL, Calhoun VD, Castellanos FX, Chen A, Chen B, Chen J, Chen X, Colcombe SJ, Courtney W, Craddock RC, Di Martino A, Dong HM, Fu X, Gong Q, Gorgolewski KJ, Han Y, He Y, He Y, Ho E, Holmes A, Hou XH, Huckins J, Jiang T, Jiang Y, Kelley W, Kelly C, King M, LaConte SM, Lainhart JE, Lei X, Li HJ, Li K, Li K, Lin Q, Liu D, Liu J, Liu X, Liu Y, Lu G, Lu J, Luna B, Luo J, Lurie D, Mao Y, Margulies DS, Mayer AR, Meindl T, Meyerand ME, Nan W, Nielsen JA, O'Connor D, Paulsen D, Prabhakaran V, Qi Z, Qiu J, Shao C, Shehzad Z, Tang W, Villringer A, Wang H, Wang K, Wei D, Wei GX, Weng XC, Wu X, Xu T, Yang N, Yang Z, Zang YF, Zhang L, Zhang Q, Zhang Z, Zhang Z, Zhao K, Zhen Z, Zhou Y, Zhu XT, Milham MP. An open science resource for establishing reliability and reproducibility in functional connectomics. Sci Data 2014; 1:140049.

[8] Jiang L, Zuo XN. Regional Homogeneity: A Multimodal, Multiscale Neuroimaging Marker of the Human Connectome. Neuroscientist 2016; 22(5):486-505.

[9] Hu Y, Wang J, Li C, Wang YS, Yang Z, Zuo XN. Segregation between the parietal memory network and the default mode network: effects of spatial smoothing and model order in ICA. Sci Bull 2016;61(24):1844-54.

[10] Gao L, Wu X, Zhang J, Chan P, Wu T. Brain activity in Parkinson's disease patients with mild cognitive impairment. Sci Bull 2016; 61(24):1876-83.

[11] Liao W, Wang J, Xu T, Zhang Z, Ji GJ, Xu Q, Wang Z, Yang F, Zuo XN, Qiu A, Zang YF, Lu G, Chen H. Altered relationship between thickness and intrinsic activity amplitude in generalized tonic-clonic seizures. Sci Bull 2016; 61(24):1865-75.

[12] Matthews PM, Hampshire A. Clinical Concepts Emerging from fMRI Functional Connectomics. Neuron 2016; 91(3):511-28.